DOCUMENT RESUME

ED 202 873 TM 810 292

AUTHOR Haladyna, Tom
TITLE A Common Metric for Integrating Research Findings.
PUB DATE [81]
NOTE 21p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Correlation; *Research Methodology; *Research Problems; *Statistical Analysis
IDENTIFIERS *Meta Analysis; *Percentage of Accounted Variance

ABSTRACT

The choice of a common metric for the meta-analysis (quantitative synthesis) of correlational and experimental research studies is presented and justified. First, a background for the problem of identifying a common metric is presented. Second, the percentage of accounted variance (PAV) is described as the metric of choice, and reasons are given for this choice. Third, two problems relating to the PAV are presented, and their solutions discussed. Finally, an optional procedure for meta-analysis is outlined which incorporates the PAV and the solutions to these problems. The characteristics of PAV are: (1) experimental or correlational results involving two variables, one an independent variable and the other a dependent variable, can be transformed into a PAV metric; (2) the metric ranges on a scale from zero to one; (3) PAV is statistically and computationally simple, and there are several statistical frameworks in which it can be computed; (4) the metric is the percentage of variance accounted in the dependent variable by the independent variable; (5) PAV can be recovered from most statistical results found in research reports; and (6) PAV is a meaningful statistic which can be generally applied to any bivariate relationship. (Author/RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
T. Haladyna
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
■ This document has been reproduced as received from the person or organization originating it.
□ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED202873

A Common Metric for Integrating Research Findings

Tom Haladyna
Teaching Research Division
Oregon State System of Higher Education

TM 810 292

The purpose of this paper is to present and justify the choice of a common metric for the meta-analysis (quantitative synthesis) of correlational and experimental research studies. The audience to which this paper is addressed are those who are actively involved in meta-analysis, as well as those who are interested in the methodology of meta-analysis.

First, a background for the problem of identifying a common metric is presented. Second, the metric of choice is described, as are the reasons for this choice. Third, two problems relating to this metric are presented, and their solutions discussed. Finally, an optional procedure for meta-analysis is outlined which incorporates this common metric and the solutions to these problems.

Background

Meta-analysis is the statistical analysis of a set of statistical analyses, focusing on the relationship between a specific dependent variable and either (a) a single independent variable or (b) a set of independent variables. The term "relationship" is important here, because what is implied in meta-analysis is not a causal analysis but merely a description of a trend over a related set of studies, which examine common variables.

The objective of meta-analysis is to obtain an estimate of the degree of relationship between a dependent and an independent variable across a large and representative sample of studies. As Glass (1979, p. 12) describes it: "One ought to reduce the various findings to a common metric, insofar as possible, and then one ought to study the pattern of these findings as

they range over many characteristics of the study themselves." Light (1979) expanded upon this thought when he asserted that we expect variation in these replications and as an additional step we need to explain why these results varied.

Since the introduction of meta-analysis by Glass (1976), there has been a rapid growth in the number of meta-analyses conducted (e.g., the effectiveness of psychotherapy, Smith, Glass, & Miller, 1980; effects of systematic instruction on learning, Kulik, Kulik, & Cohen, 1979; relationship of student, teacher, and learning environment variables to attitudes toward the sciences, Haladyna, Olsen, & Shaughnessy, 1980; relationship of motivation to school learning, Uguroglu & Walberg, 1979).

Despite the increasing use of meta-analysis, many methodological problems persist (Glass, 1978; Jackson, 1980), and there have been proposed solutions to some of these problems (Barton & Glass, 1979, McGaw & Glass, 1980).

The Choice of a Common Metric

One persisting problem is that of a common metric. That is, a statistical measure is needed that can be applied effectively to both experimental and descriptive studies. The most commonly used metric is effect size (ES), which is simply the normalized difference between two group means. The formula for ES is:

$$ES = \frac{\overline{X}_E - \overline{X}_C}{S_X}, \text{ where}$$

the means $\overline{X}_E$ and $\overline{X}_C$ represent experimental groups and

S_X is the within groups' standard deviation.

One of the most appealing characteristics of ES is its ease of interpretation; another is its ease of computation. However, as Glass (1978) and Jackson

(1980) have noted, a number of limitations exist with the use of ES. One of these limitations is that ES applies only to group differences which arise from experimental studies. Another problem is that when the contrasting groups have unequal variances, our choice of a standard deviation yields different ES's and, consequently, different interpretations. Another problem is that the metric scale used in one study is unlike that used in another. Also, the measures serving to represent our dependent variable often differ from study to study. A good illustration of this is the meta-analysis of science attitudes conducted by Haladyna and Shaughnessy (1980), where a variety of distinctly different definitions of attitude toward science existed, some of which were quite conceptually independent of the others. Finally, research reports do not consistently report statistics that can be aggregated using ES.

While McGaw & Glass (1980) have addressed the problem of a common metric for ES within the framework of complex experimental designs, little progress has been made on meta-analyses which span experimental and correlational studies. Jackson (1980, p. 456) conveys this attitude toward meta-analysis when he says:

> ...there are better methods for doing it (meta-analysis) than are currently being applied in most reviews.

The definitive work on the methodology of meta-analysis was done by Glass (1978). In that paper, ES is given major attention, as it has been successfully applied to a number of review topics. A methodology for correlational studies also exists, in parallel to those established for experimental studies, but there have not been any discussions of a general methodology for meta-analyses which spans both experimental and correlational studies, although Glass (1978, pp. 372-375) does devote some attention to

the general topic of correlational studies in meta-analysis. It is the primary objective in this paper to extend this work to generalize to all data analyses aggregations.

Characteristics of the Proposed Metric

A metric which is common to both experimental and correlational studies is the percentage of accounted variance (PAV). The characteristics of PAV are viewed as quite desirable to meta-analysts and are presented below:

1. First, any experimental or correlational result involving two variables, one an independent variable and the other a dependent variable, can be transformed into a PAV metric. Therefore, one does not have to compute ES for one set of studies and product-moment correlations -- (r), r^2, or r-z-transformations -- for the other set of studies.

2. The metric ranges on a scale from zero to one. Most statistical results will fall in the lower range of the scale, from zero to .10, the latter representing a product-moment relationship of about .32.

3. PAV is statistically and computationally simple, and there are several statistical frameworks in which it can be computed. For instance, one could approach the computation of PAV using the general linear model (ordinary least-squares regression analysis), but more often any beginning textbook on statistics which includes such topics as correlational methods and analysis of variance will be sufficient as a reference.

4. The metric is easy to interpret, taking the familiar form observed with the use of the squared multiple correlation coefficient (R^2), which is the coefficient of determination (Guilford, 1965). Literally, it is the percentage of variance accounted in the dependent variable by the independent variable.

5. PAV can be recovered from most statistical results found in research reports. As Glass (1978) and others have observed, there is little consistency in research reports. In fact, there is a beleaguering variety of statistical methods and omissions of crucial data, such as means and standard deviations, in many reports. Despite the frustrations of aggregating the seemingly unaggregatable, there are methods of recovering statistics that lead to the estimation of PAV.

6. PAV is a meaningful statistic which can be generally applied to any bivariate relationship.

The essence of the PAV metric. In aggregating statistical results across studies, the task begins with estimating PAV for a bivariate relationship between a single dependent and a single independent variable. Any bivariate relationship can be reduced to any one of the seven conditions shown in Table 1. When the two variables of interest are dichotomous, the product-moment technique is used (known as the phi coefficient).

When the dependent variable is dichotomous and the independent variable is continuous, the reported statistic might be a point-biserial correlation or a t-statistic. The point-biserial correlation is the product-moment correlation when one of the variables is dichotomous and the other is continuous. There is a simple formula for transforming a t-statistic to the point-biserial shown in Table 1. The squared point-biserial is PAV.

When both variables are polychotomous, and the categories have an underlying order (which is usually the case), one might see a chi-square reported in a study. In that instance, a formula is provided in Table 1 for estimating the product-moment relationship. If raw data are available in the report, one might treat the data as continuous and compute the actual product-moment coefficient.

Table 1

Type of Dependent Variable	Type of Independent Variable	Common Solution[1]	Per Cent of Accounted Variance (PAV)
Dichotomous	Dichotomous	phi coefficient (ϕ)	ϕ^2
Dichotomous	Continuous	point biserial (r_{pbi}) or t-test $r_{pbi} = \sqrt{\frac{t^2}{t^2 + N_1 + N_2 - 2}}$	r_{pbi}^2
Dichotomous	Polychotomous	multiple correlation (R)	R^2
		eta coefficient (η)	η^2
Polychotomous	Polychotomous	$r_{xy} = \sqrt{\frac{\chi^2}{\chi^2 + N}}$	r^2
Continuous	Dichotomous	point biserial, t-test	r_{pbi}^2
Continuous	Continuous	product moment (r) for a new linear relationship	r^2
Continuous	Continuous	eta squared for a new linear relationship	η^2

[1] All of these problems can be solved within the framework of ordinary least-squares regression analysis.

When the dependent variable is dichotomous and the independent variable is polychotomous, the multiple correlation ($\underline{R}$) is often reported, and $\underline{R}^2$ is our PAV. In this situation, we can also compute an eta coefficient which leads to the identical result.

When the dependent variable is continuous and the independent variable is dichotomous, the point-biserial or $\underline{t}$-statistic may be reported. The transformation to PAV is the same as reported when the dependent and independent variables were dichotomous and continuous, as shown in Table 1.

When both variables are continuous, two possible statistics may be reported. The first is the product-moment correlation, where $\underline{r}^2$ is PAV, and the second is the eta coefficient, where $\underline{\eta}^2$ is PAV.

The entries in Table 1 provide the majority of statistical results reported in research papers which lead to PAV estimates. However, there are instances where data are not fully reported or statistics are referred to in a manner in which the PAV is not easily estimated. Some of these instances and the manner in which PAV is recovered may be found in the appendix. The approaches suggested are not fully comprehensive of all possibilities, but offer opportunities for obtaining PAV estimates in these instances. Some creativity and algebra are often needed with cases that arise that are not covered in the more conventional cases shown in Table 1 or the less conventional cases given in the Appendix.

The relationship of PAV to ES. While ES has been given major attention in many meta-analytical reviews, the advocacy of PAV as a more useful and general metric should not be viewed as departing greatly from prior practice. In fact, there is an isomorphic relationship between ES and PAV estimates. The relationships between these two common metrics, when the groups are compared, are expressed below:

$$ES = \bar{X}_a - \bar{X}_b \sqrt{\frac{PAV(N - 1)}{SS_B}}$$

$$PAV = \frac{SS_B ES^2}{(\bar{X}_a - \bar{X}_b)^2 (N - 1)}$$

where

$\bar{X}_a$ and $\bar{X}_b$ are group means,

SS_B is the between sum of squares, i.e.,

$n_a(\bar{X}_a - \bar{\bar{X}})^2 + n_b(\bar{X}_b - \bar{\bar{X}})^2$, and

N is the total sample size.

The above equivalencies for ES and PAV are expressed in familiar terms and refer to two-group comparisons for which a t-test is normally run. Within the framework of one-way or factorial analyses of variance, the equivalencies are more complex to express.

An inspection of the two expressions above indicates that the two are exchangeable, but PAV is favored for the reasons provided earlier in this paper, the most important of which is that PAV is a more general measure of effect size and, therefore, applicable to the wide variety of statistical results found in research reports.

The discussion now turns to two problems which are related to the choice and use of the PAV metric in meta-analysis. The first problem has to do with the statistical null hypothesis and the estimation of the PAV, and the second problem has to do with the unit of analysis and size and interpretation of PAV.

The Null Hypothesis and PAV (Capitalizing on Small Samples)

In any statistical test, the null hypothesis takes the general form that there is no difference in groups being tested or no relationship between the

independent and dependent variables. In traditional statistical parlance, these are generally stated:

H_o: $\mu_1 = \mu_2 = \mu_3 \cdots \mu_n$

H_o: $\rho = 0$

As we know from the lengthy debates that have been conducted recently (Cohen & Hyman, 1979, 1980; Hsu, 1980), it is reasonable to evaluate results in terms of practical significance following a test for statistical significance. However, it is also common practice to take the results of a statistical analysis, where the null hypothesis was accepted, and aggregate these results as if the resulting statistics were accurate population estimates.

What seems to be a more reasonable practice is to test for the null hypothesis. If the statistic is not significant, we then accept the null hypothesis that $\mu_1 = \mu_2 = \mu_3 \ldots \mu_n$ or $\rho = 0$ and PAV = 0, and aggregate these results in the magnitude that our decision rule dictated.

A case in point can illustrate this. In analyzing the relationship between X and Y across a multitude of studies, we come across a correlation of .30, which is non-significant at .05, for a sample size of 35. Do we treat this result as $\rho = .30$ or $\rho = 0$? Our statistical decision rule was that $\rho = 0$, therefore, $\rho \neq r$ and PAV = 0.

Cohen (1977, p. 10) states:

> ...the ES can itself be treated as a parameter which takes the value zero when the null hypothesis is true and any other specific none zero value when the null hypothesis false....

Hsu (1980, p. 8) adds:

> Failure to find statistical significance in a study indicates that the sample ES is consistent with the hypothesis that the actual magnitude of the effect (i.e., the actual ES) = 0.

Therefore, it would seem reasonable that meta-analysts should use the statistical decision rules to determine whether or not PAV equals zero and abide by that decision when aggregating results. To include a correlation of .30, as in the example previously provided, when it is not statistically significant, introduces a serious source of error into our results.

The Unit of Analysis in Meta-Analysis

Another persistent problem in data aggregation is the unit of analysis upon which the statistical result is based. Recently, there have been a growing number of studies where the units of analysis are no longer individuals, but classrooms or schools (e.g., Rutter, Maughan, Mortimore, & Ouston, 1979; and Brookover, Beady, Flood, Schweitzer, & Wisenbaker, 1979).

What is not surprising, according to Sirotnik (1980) and Light (1979), is the fact that results vary as a direct function of the unit of analysis. As Sirotnik explains it, there are two aspects to this problem: statistical and interpretive. The statistical aspect involves the fact that one set of results based on a particular unit of analysis may be quite different than another set of results based on another unit of analysis. The nature of this statistical difference has been widely discussed in recent writings (e.g., Burstein, 1975; Knapp, 1977, Light, 1979), and Light provides a simple-to-compute formula for the estimation of PAV for different units of analysis.

The second aspect of this problem deals with how we interpret these results. A result at one unit may have one kind of meaning, but at another level the result may be interpreted differently. In other words, we need to address the rationale behind the unit of analysis before aggregating it with other units of analysis.

At issue in meta-analysis is that the unit of analysis should always be considered for these two reasons, statistical differences that occur with the same data at different units of analysis, and the ensuing difference in interpretations. To aggregate PAV's from these differing units of analysis distorts the estimation of PAV over the sample of studies being synthesized. As in the case of the use of the statistical null hypothesis, there is a possible other source of bias in the estimation of the average PAV.

A good illustration of this problem of unit of analysis was demonstrated in a series of studies of attitudes toward mathematics (Shaughnessy, Haladyna, & Shaughnessy, 1980, 1981). In this study, nearly 2000 fourth, seventh, and ninth graders were examined with respect to relationships between attitudes toward mathematics and student, teacher, and learning environment variables for two units of analysis. Some results of these respective units of analysis are given below:

Grade 7

	Correlation with Individual Math Attitudes (n=764)		Correlation with Class Mean Attitudes (N=34)	
Independent Variables	r	PAV	r	PAV
Fatalism	-.43	18	-.60	36
Self-confidence	.35	12	.41	17
Teacher Quality	.48	23	.78	61
Satisfaction	.33	11	.72	52
Class Environment	.33	11	.55	30
Disorganization	.32	10	.56	31

As shown above, the sizes of the relationships and PAV's vary systematically as a function of the unit of analysis. Light (1979, p. 10) provides a formula which illustrates the dependency of these PAV's upon the intraclass

correlation of the dependent variable with the independent variable (the classes). When classrooms differ substantially, the ratio of between and within variances yields a sizable F which results in large discrepancies in PAV's for individual and class correlations. For example, the above data would suggest a large variation in the class means of teacher quality ($F = 11.2$, df 33,730) indicating that teacher quality has a stronger influence on student attitudes toward math at the class level than at the individual level. The case for this occurring with student fatalism is not as strong.

In conclusion, relationships between a dependent and independent variable have been shown to systematically vary as a function of the unit of analysis. Meta-analysts need to be aware of this phenomenon and deal with it at statistical and interpretive levels. First, there is a statistical difference which manifests itself in two discrepant sets of results. One should not aggregate PAV's for these differing units of analysis. Second, there is often a difference in interpretations. Therefore, one needs to acknowledge both of these levels of the problem and address their meta-analyses to these levels. Operationally, this means two separate sets of analyses for the same relationships, one for each unit of analysis.

An Optimal Procedure for Conducting a Meta-Analysis

The foregoing discussion has emphasized the use of a common metric, PAV, for aggregating research findings, acknowledging the differences arising from the units of analysis. Further, it has been advocated that a statistical null hypothesis be first tested to determine its truthfulness and the decision be used to determine the PAV. When the null hypothesis is accepted, PAV = 0.

Now it seems appropriate to prescribe a set of procedures for the conduct of meta-analysis which includes some of these principles and practices. These optimal procedures should include those given here:

1. Identify the variables by which we describe each statistical result. If meta-analysis really involves the study of patterns relating to PAV, then the identification of these variables is critical. These variables may include years studies were reported, grade levels, gender of students (or subjects), sizes of samples, sources of publication, types of independent variables, reliability of data, or any other variable considered salient to the problem under study.

2. Test for the null hypothesis. When the null is accepted, PAV = 0. Aggregate accordingly.

3. Insure that when the unit of analysis varies, results are tabulated separately accordingly. Maintain an integrity in interpretation, asking the basic question, "What am I looking for?" or "What does this result mean?" as Sirotnik (1980) has suggested.

4. Compute the PAV for each result. Use the standard techniques reported in Table 1, or, when data are more obscurely reported, use the Appendix[1] or develop alternative recovery procedures (in consultation with your friendly statistician).

5. Table results in a meaningful display. One such method which has a great deal of clarity is the stem-and-leaf method prescribed by Tukey (1977) and successfully employed in such meta-analyses as the one by Uguroglu and Walberg (1979).

To overcome the problem of over- or underrepresentation of results from any one study, use such procedures as the jackknife method (which reduces the tendency for a single study with abundant results to overrepresent itself).

---

[1] The Appendix provides some techniques for recovering PAV's from research reports where the more conventional procedures fail.

These procedures are well described in Glass (1978). A good illustration of this problem was in a meta-analysis on math attitudes by Shaughnessy, Haladyna, Olsen, and Shaughnessy (1980), where nearly half the reported values in the meta-analysis came from a single study!

6. Employ appropriate analytical methods, such as regression analyses, to determine if any of these variables identified in #1 are systematically related to PAV.

Since meta-analysis is the quantitative synthesis of research reports, results can serve several objectives. If we view such reviews as synthesizers of knowledge, they should lead to policy or research decisions which will enable us more effectively to plan programs or to plan further research. There may be other purposes for research integration, but, for the most part, these kinds of reviews have become useful summaries of research as an aid to decision making.

The choice of a common metric and the presentation and suggested solutions to two problems accompanying the use of this metric are intended to both simplify and properly rectify practices of meta-analysis. By no means does this resolve the abundant other issues that other writers have identified as problems with the procedures of meta-analysis.

References

Barton, M. A., & Glass, G. V. Integrating studies that have quantitative independent variables. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 1979.

Brookover, W., Beady, C., Flood, P., Schweitzer, J., & Wisenbaker, J. School social systems and student achievement: Schools can make a difference. Brooklyn, N.Y.: J. F. Bergin Publishers, Inc., 1979.

Burstein, L. The role of levels of analysis in the specification of educational effects. Los Angeles: Graduate School of Education, UCLA, 1978.

Cohen, J. Statistical power analysis for behavioral sciences. New York: Academic Press, 1977.

Cohen, S. A., & Hyman, J. S. How come so many hypotheses in educational research are supported? Educational Researcher, 1979, 8, 12-16.

Cohen, S. A., & Hyman, J. S. Where would we be without alpha? A new convention for upgrading scientific rigor in educational inquiry. Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

Glass, G. V. Primary, secondary and meta-analysis of research. Educational Researcher, 1976, 5(10), 3-8.

Glass, G. V. Integrating findings: The meta-analysis of research. Review of Research in Education, 1978, 5, 351-379.

Glass, G. V. Policy for the unpredictable (uncertainty research and policy). Educational Researcher, 1979, 8(9), 12-14.

Guilford, J. P. Fundamental statistics in psychology and education. New York: McGraw-Hill, 1965.

Haladyna, T., Olsen, R., & Shaughnessy, J. Relations of student, teacher, and learning environment variables to attitudes toward science. Manuscript to be submitted for publication, 1980.

Haladyna, T., & Shaughnessy, J. Attitudes toward science: A quantitative synthesis. Manuscript submitted for publication, 1980.

Hsu, L. M. On why many hypotheses in educational research are supported and on the interpretation of sample effect sizes: A comment. Educational Researcher, 1980, 9(5), 6-8.

Jackson, G. B. Methods for integrative reviews. Review of Educational Research, 1980, 50(3), 438-460.

Knapp, T. R. The unit-of-analysis problem in applications of simple correlation analysis to educational research. Journal of Educational Statistics, 1977, 2, 171-186.

Kulik, J. A., Kulik, C., & Cohen, P. A. A meta-analysis of outcome studies of Keller's personalized system of instruction. American Psychologist, 1979, 34(4), 307-318.

Light, R. J. Capitalizing on variation: How conflicting research findings can be helpful for policy. Educational Researcher, 1979, 8(9), 7-14.

McGaw, B., & Glass, G. V. Choice of the metric for effect size in meta-analysis. American Educational Research Journal, 1980, 17(3), 325-337.

Rutter, M., Maughan, B., Mortimore, P., & Ouston, J. Fifteen thousand hours: Secondary schools and their effects on children. Cambridge, Mass: Harvard University Press, 1979.

Shaughnessy, J., Haladyna, T., Olsen, R., & Shaughnessy J. M. Student attitudes toward math: A meta-analysis. Manuscript submitted for publication, 1980.

Shaughnessy, J., Haladyna, T., & Shaughnessy, J. M. Determinants of class attitudes toward math. Manuscript submitted for publication, 1980.

Shaughnessy, J., Haladyna, T., & Shaughnessy, J. M. Relations of student, teacher, and learning environment variables to attitudes toward math. Paper to be presented at the annual meeting of the American Educational Research Association, Los Angeles, April 1981.

Sirotnik, K. A. Psychometric implications of the unit-of-analysis problem (with examples from the measurement of organizational climate). Journal of Educational Measurement, 1980, 17(4), 245-282.

Smith, M. L., Glass, G. V., & Miller, T. I. Benefits of psychotherapy. Baltimore: Johns Hopkins University Press, 1980.

Tukey, J. W. Exploratory data analysis. Reading, Mass.: Addison-Wesley Publishing Company, 1977.

Uguroglu, M. E., & Walberg, H. J. Motivation and achievement: A quantitative synthesis. American Educational Research Journal, 1979, 16(4), 375-390.

Appendix

Most research reports contain information about the statistical analyses which allows for a transformation of results into the percentage of accounted variance (PAV). However, it is recognized that research reports will occasionally be somewhat obscure in the presentation of statistical results, and less conventional means are needed to recover PAV estimates. This appendix is devoted to expanding the set of procedures presented in the paper to accommodate those instances where incomplete or inadequate statistical results are presented and some novel means are needed to estimate PAV.

Correlational Results

Correlational results not developed in the paper are discussed here. For the most part, most correlation techniques are a form of the product-moment correlation or the eta coefficient where the simple squared coefficient is equal to PAV. Two rank-order techniques, the Spearman coefficient and tau, have been neglected. Each is based on ordinal data and each cannot be directly translated to a product-moment estimate. However, both are underestimates of the product-moment coefficient. That is to say, given two continuously measured variables which are reduced to ordinal data, the comparison between a product-moment coefficient and a Spearman or tau coefficient would show that the latter two are underestimations of the former. Therefore, if either of these rank-order correlations are used, we must accept the fact that they are probably negatively biased estimates.

Group Differences

The second case covers group differences where most of the data analysis problems exist. The following material is an explication of problems actually encountered in several meta-analyses where standard statistical

results were not clearly presented and the recovery of the PAV was made difficult. In each of these instances, a simple relationship was employed:

$$PAV = \frac{\text{between sums of squares}}{\text{total sums of squares}}$$

This expression is the equivalent of the squared eta coefficient. The major objective is, then, to identify or compute these two sums of squares and then take the ratio of these to estimate PAV.

Using conventional notation, let

ΣX^2 = the sum of the squares of the scores;

N = the total sample size;

n_g = the number of groups;

S = the standard deviation for the total sample;

SS_b = the between sum of squares;

SS_t = the total sum of squares;

MS_b = the between mean square $[SS_b(n_g - 1)]$;

MS_t = the total mean square $(SS_t / n - 1)$ or S^2;

F = the F-ratio;

t = the t-statistic $(t^2 = F)$;

$\overline{\overline{X}}$ = the grand mean; and

$\overline{X}_i$ = the mean for group i, i - 1...n_g.

Case One: Only means and standard deviations are given.

Compute the sum of the scores $(\Sigma X = N\overline{\overline{X}})$ and the sum of the squares of the scores $[\Sigma X^2 = (N - 1)S^2 + \frac{(\Sigma X)^2}{N}]$.

Combine these sums and use the formula for the standard deviation to compute the sum of squares for the total: $SS_t = \Sigma X^2 - \frac{(\Sigma X)^2}{N}$.

Compute the sum of squares using the formula below:

$$SS_b = \sum_{i=1}^{n} n_i \overline{X}_i^2 - N\overline{\overline{X}}.$$

Take the ratio of the two sums of squares to get PAV.

Case Two: Given an F-ratio or t-statistic, and the sample size and group means,

1. Compute the sum of squares between (as shown in Case One).
2. If two groups, $t^2 = F$.
3. Solve for sum of squares total:

$$SS_t = \frac{SS_b[(N - n_g) + F(n_g - 1)]}{F(n_g - 1)}$$

Case Three: Given an F-ratio and N,

1. Find the sum of squares for the total:

$$SS_t = S^2(N - 1).$$

2. $$SS_b = \frac{F(SS_t)(n_g -)}{(N - n_g) + F(n_g - 1)}$$

Case Four: F-ratio, MS_b, and MS_w are given.

$$SS_b = (MS_b)(n_g - 1)$$

$$SS_w = (MS_w)(N - n_g + 1)$$

$$SS_t = SS_b + SS_w$$

Postscript

The few methods displayed here only scratch the surface of possibilities. What this appendix should emphasize is that a certain amount of resourcefulness and algebra is needed to recover these estimates from obscurely reported data. But the effort results in a more adequate treatment and aggregation of results.